Strategic Interdependent Security Systems

Thomas A. Weber
École Polytechnique Fédérale de Lausanne
thomas.weber@epfl.ch

Abstract

We develop a model of investments in interdependent security systems in the presence of a strategizing opponent, who attempts to infiltrate at least one of the systems with a harmful device. In the first stage, a finite number of targets decide about their respective investments in direct and indirect security systems by choosing appropriate detection probabilities minimizing expected losses. In the second stage, infiltration of a target's security system with the harmful device is directed by the agent from the outside and may occur either directly by a target's failing to detect the device at its entry level or indirectly by first passing undetected through the other target's entry screening and subsequently clearing the cross-screening device.

1 Introduction

Recent history has shown that private organizations (targets) as well as public institutions may fall victim to malevolent attacks from the outside directed at maximizing the losses to their targets and possibly society at large. As a consequence of the intertwined relationships between many of the targets and the strategizing attitude of the malevolent outside force plotting an attack, private investments in security systems are highly *interdependent*, a notion which will be made more precise below. Our paper focuses on these interdependencies and their consequences from the perspective of both a private target and a welfare-maximizing social planner, of which the latter is in a position to change the ground rules through regulation and public deterrence.

We summarize by "security" all measures that aim at protecting a target's facilities and at providing freedom from danger or anxiety to individuals that could be directly or indirectly affected by an attack. An attack may come in the form of a terrorist operation using, for example, an explosive device to sabotage productive facilities and cause individual harm, or it may manifest itself in other malevolent activities, such as the theft of valuable assets, manipulation or jamming of business activities, resulting in loss of control, or simply diversion of decision makers' attention. In assessing the overall damage caused by a successful attack one needs to consider the affected target's foregone payoffs and a monetary equivalent for the harm or inconvenience to individuals both within and outside the target. Overall losses are thus incurred partially by the attacked target and its employees (the *private losses*) and partially by the rest of society (the *public losses*). In other words, the impact of a successful attack consists in its immediate effect on the targeted organization *and* its repercussions (or negative externalities) on the rest of society.

For concreteness and to streamline our analysis, we assume that the threat of an attack emanates from a single pernicious agent whose aim it is to maximize the extent of damage to any targeted organization or, possibly, to society at large. Multiple pernicious agents and the possibility of multiple attacks in a dynamic setting are interesting topics for future research and are not considered here. We further suppose that the provision of security systems is costly and decisions about their acquisition are made autonomously by the different targets. In this context, by a "security system" we mean all the various elements within a target that provide access control and safety monitoring so as to ensure early detection and thus attack prevention.

Given the pernicious agent's goal to maximize the damage caused by his attack (provided that it is individually rational for him to carry it out), the targets' respective investments in security systems are *interdependent* in two distinct ways. First, a lack of security investment – all else equal – renders a target more prone to an attack, since the pernicious agent trades off expected losses (i.e., his "gains") against the expected disutility of being detected and punished. As an interesting consequence of the agent's optimizing behavior, the targets may be in direct competition with each other to provide the highest security level in order to divert an inevitable attack. We denote these interdependencies caused through *target substitution* on the part of the pernicious agent as "first-order" interdependencies. Second, the targets' security systems themselves may be interdependent as a result of the targets' interconnectedness. For instance, the targets may be using

URI: https://hdl.handle.net/10125/80193
978-0-9981331-5-7

a shared resource such as a common infrastructure. As a result, the agent might carry out the attack indirectly by passing first undetected through another unit's access control system and then use private links *between* different units to reach the ultimate target. To take account of such "second-order" interdependencies we distinguish between "direct" and "indirect" security systems, whereby the former provides safety against *direct attacks* while the latter shields against *indirect attacks*. We furthermore assume that for each target k there exists a simple statistic in the form of a *detection probability* p_k, which completely characterizes the overall quality of its security system, both in terms of its robustness against a direct attack (p_k) and against an indirect attack (q_k) being carried out by passing first through another organization.[1]

In addition to the targets' security systems, complementary protective measures may be provided by executive and judicial branches of societal institutions such as the police force, national security agencies, the military, special task forces or, to some extent the court and legal system. In mobilizing public means for the provision of deterrence a social planner needs to balance the private and public incentives for preventing harmful attacks on targets. For this he needs to consider two countervailing effects. First, as alluded to earlier, the targets generally do not bear the full social cost of an attack and may therefore tend to privately underinvest in security systems. Second, the intertarget competition resulting from the perceived target substitution effect may increase security levels above the socially optimal level. The planner in turn uses a range of measures to respond to these effects and at least partially re-internalize the potential negative externality resulting from the threat of attacks to the private targets, increasing their security spending. These measures may include regulation in the form of safety standards and public deterrence in the form of prosecution, a credible threat of retaliation, or other appropriate punishments.

The model developed here differs from prior models in a number of ways: we examine a fully strategic setting, consider regulatory action, and admit incomplete information on the targets' respective loss distributions.

1.1 Literature

There is ample literature concerning "security problems," broadly encompassing a class of problems that targets and possibly individuals can face, with the following two features:[2]

a. *Existence of an outside threat:* there is an outside force that can cause damage to the targets and individuals. The behavior of the outside force can be *either* nonstrategic (e.g., a natural disaster), so that it is independent of what action any target takes, *or* strategic (e.g., a burglar or terrorist), so that it may be directly influenced by the targets' behavior.[3]

b. *Payoff interdependency:* the behavior of one target in the face of potential losses impacts the subjective loss distribution of other targets and consequently their self-interested behavior. These payoff dependencies can be of first-order (target substitution) and/or second-order (interconnectedness), as explained earlier.

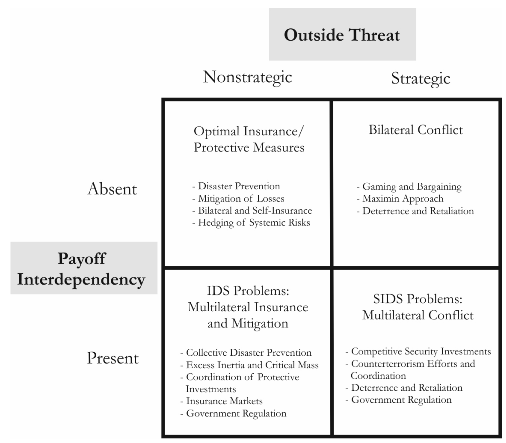

Figure 1: Classification of Security Problems.

We denote the class of problems exhibiting feature b. with at least first-order interdependencies as *interdependent security* (IDS) problems. Payoffs in IDS problems depend on both the nature of the outside threat *and* the payoff externalities between agents. In this paper we focus on *strategic* IDS (or SIDS) problems, in which the outside threat is assumed to be self-interested and rational. To clarify the main differences in our approach, that admits both first- and second-order payoff interdependencies, we now briefly review the different classes of security problems; cf. Fig. 1.

Optimal Insurance and Protective Measures. If the outside threat is nonstrategic and externalities between different agents are absent, then the overall security problem decomposes into a number of independent single-person decision making problems. Each decision maker must choose his optimal insurance coverage given his set of beliefs about the loss distribution

[1] As an example, the probability of a successful indirect attack on target 1 through target 2 is $(1-p_2)(1-q_1)$, cf. Fig. 2.

[2] Enders and Sandler [5] provide an interesting overview of literature on security problems specifically related to terrorism.

[3] In both cases there may be considerable uncertainty associated with the magnitude of a potential loss. Furthermore, the prior beliefs about the loss distributions may vary across individuals.

implied by the external threat [16]. This analysis carries over to the modeling setting in which an inherently strategic outside threat (such as terrorism) is approached from a probabilistic single-period point of view. Using an open-loop (i.e., non-equilibrium) approach it is possible to consider protective measures in considerable detail using, for example, a decision analysis framework by allocating resources across the threatened system from the viewpoint of a single decision maker [12].

Bilateral Conflict. If a single target faces the threat of a single pernicious agent who acts strategically according to rational preferences, then in the resulting zero-sum game it is often in the target's best interest to minimize its worst-case payoffs which at the same time minimizes the agent's incentives for an attack. If the losses from an attack are reversible or deferred, the final payoffs may be contingent on the outcome of a negotiation between the target and the pernicious agent. As an example, Selten [15] examines a simple model of kidnapping and concludes based on a model with perfect information that it is generally not in the kidnapper's best interest to carry out a threat of killing his victim. Nevertheless, an (irrational and exogenous) commitment to a *mixed* strategy (the kidnapper kills his victim only sometimes) turns out to provide a higher payoff.[4] A more systemic point of view is developed by Hirshleifer [7], who examines an equilibrium model for the resource allocation of two neighboring economies between warfare ("appropriative effort") and productive activities ("contestable productive effort").

IDS Problems: Multilateral Insurance and Mitigation. [6] and [10] consider IDS problems in a variety of contexts, such as the choice of individuals in vaccinating or not against an infectious disease. Even though their analysis captures the payoff interdependencies resulting from shifts in the threat when other agents make their vaccination decisions, the outside threat remains non-strategic. In other words, the probability of a direct infection is fixed and independent of the protective actions of other agents, so that rational target substitution (a natural consequence of strategic outside behavior) cannot occur. Under these conditions [6] shows that security investment may be subject to excess inertia in the sense that if nobody has invested in security, then incentives to go first might be minimal. On the other hand, if many players adopt a high security standard, there may be the possibility for free-riding which endogenously limits the security investment, generating multiple equilibria and thus a need for coordination.

SIDS Problems: Multilateral Conflict. In situations where payoff externalities between the different players are present and the external threat exhibits strategic behavior, the players' actions will be taken both in competition with each other and in anticipation of the pernicious agent's equilibrium actions. [9] uses simulation to approximate an equilibrium model, but the "avoidance" dynamics they obtain are a consequence of rather specific assumptions on the way substitution between targets occur. [14] examines an IDS problem where the target substitution effect is captured by a first-order stochastically dominant shift of two countries' prior beliefs about their being attacked, as a result of their respective investment in deterrence. The focus is on symmetric equilibria with perfect information, where the countries are identical and thus the pernicious agent's payoffs are the same from a successful attack on either one of the countries. As pointed out earlier we do admit private information about the targets' expected losses and consider the interesting case where targets are interconnected, whence the pernicious agent has different modes of attack at his disposal, direct and indirect. It turns out that in equilibrium the agent has no strong preference for an indirect attack. Indeed, if the protection against one mode increases, the pernicious agent substitutes away from that mode to either a different mode of attack or a different target altogether. [3] examines the effectiveness of policy interventions responding to terrorism and finds empirical evidence for the substitution phenomenon in two out of three evaluated policies. Their notable exception is the installation of metal detectors in US airports starting January 1973 that induced a significant drop in skyjackings and (after a three-month lag) a substantial increase in non-skyjacking terrorist activity. In our analysis we do not distinguish between specific modes of attack other than direct or indirect, however substitution (or "transference" in the criminal justice literature) is a natural consequence of the pernicious agent's self-interested optimizing behavior evaluating the relative attractiveness of different targets and modes of attack. Thus, rendering one target less attractive may lead the pernicious agent to switch his strategy.

1.2 Outline

This paper examines both private and public choice in the presence of multiple payoff interdependencies between targets, pernicious agent, and a social planner. Section 2 introduces a three-stage model. Section 3 presents the pernicious agent's optimal attack policy as well as a symmetric Bayes-Nash bidding equilibrium as the solution to the last two stages of the model. Section 4 then treats the regulator's policy problem in providing the correct incentives for a social-welfare maximizing solution. Section 5 concludes.

[4] If the game is repeated, such as in hijacking incidents between governments and terrorist organizations, then a no-negotiation policy is typically not in a government's best interest [11].

2 Model

We consider a situation in which $K \geq 2$ targets (e.g., firms or public organizations), depending on their investments in security systems, are vulnerable to a deliberate attack by a pernicious agent. The agent's goal is to cause what he perceives to be the maximum expected loss by setting off a single harmful device (e.g., an explosive) in one of the targets, while taking into account the likely negative consequences for him in case an attack fails. In addition to the equilibrium behavior of both agent and targets in the associated sequential-move game we are concerned with a regulator's socially optimal choice of minimum safety standards as well as the question of what public means should be used for implementing measures directed at deterring the agent from carrying out his malevolent attack in the first place.

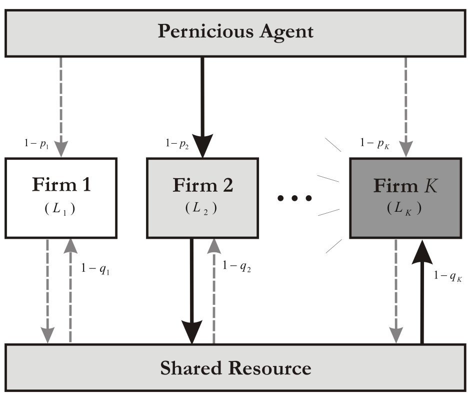

Figure 2: Model Overview.

Overview. A main feature of the model is that all targets are interconnected by a shared resource (cf. Fig. 2): an attack on target $k \in \mathcal{K} = \{1, 2, \dots, K\}$ may be conducted either *directly* by entering through its entry-level security systems (detection probability: p_k) or *indirectly* by infiltrating another target $j \neq k$ first (detection probability: p_j) and then passing the harmful device internally via the shared resource from target j to target k (detection probability: q_k). Such an attack may be *either* successful, in which case target k faces a nonnegative expected loss L_k, *or* it may be unsuccessful if the target is able to detect the harmful device, in which case the pernicious agent faces prosecution resulting in an expected nonpositive payoff of $-\delta$ for him, where $\delta \geq 0$ is a given constant. The targets' prior beliefs about the distribution of their losses are common knowledge and not a source of disagreement [1]. The pernicious agent evaluates the loss prospects prior to deciding first *if* to attack, then *whom* to attack, and eventually *how* to attack. To counteract the threat of an attack each target can independently invest in security systems to increase its direct detection and cross-check capabilities p_k and q_k, respectively. Anticipating the targets' and the agent's behavior, the regulator attempts to maximize social welfare by imposing a minimum security level (or "safety standard") ρ for entry-checks, effectively imposing the constraint $p_k \geq \rho$, $k \in \mathcal{K}$, on the targets' investment problem.[5] In addition to requiring a minimum security investment from agents, the regulator can also make an effort in prosecution and punishment after detection, effectively deterring the agent by increasing his expected *dis*utility in case of an unsuccessful attack. The regulator, the targets, and the agent move sequentially over three time periods $t \in \{0, 1, 2\}$; cf. Fig. 3. We solve for symmetric (perfect) Bayesian equilibria by using backward induction.

Agent. At time $t = 2$, the pernicious agent finds out the agents' vulnerabilities in terms of their expected losses and security systems. Thus, the agent is assumed to know more than any single target when plotting his attack, which is a conservative assumption that leads to robust security design choices for the targets. For a direct attack on target k the agent knows the expected loss L_k, the target's direct security level p_k, as well as the regulator's choice of δ and ρ. The agent's payoff from a direct attack on target k is $\varphi_k^d = L_k - p_k(L_k + \delta)$. For an indirect attack on target k the agent can choose the path of least resistance by passing through the entry systems of the target with the lowest security, $\bar{p}_{-k} = \min_{j \in \mathcal{K} \setminus \{k\}} \{p_j\}$. Our conservative assumption is that the agent is able to devise a path into the shared infrastructure at the minimum admissible detection threshold ρ which never exceeds $\bar{p}_{-k}$. This assumption provides robustness of the target's security investments against a later expansion of the target network, as the equilibrium safety investments are nonincreasing in the number of targets. It also renders the analysis tractable, since the distribution of the order statistic $\bar{p}_{-k}$ is endogenous and itself determines the equilibrium. It also makes the agent, whose payoff from an indirect attack on target k is $\varphi_k^i = L_k - (q_k + (1 - q_k)\rho)(L_k + \delta)$, somewhat more prone to attack. Based on the comparison of φ_k^d and φ_k^i the agent decides whether to attack k depending on whether the best *mode* of attack m, which could be either *direct* ($m = d$) or *indirect* ($m = i$), yields a nonnegative expected payoff $\varphi_k = \max_{m \in \{d,i\}} \{\varphi_k^m\}$ for him. These payoffs are given in Table 1. Conditional on his attack (which by itself has an opportunity cost of zero) not being successful, the agent expects a disutility of δ resulting from the blow of his cover and additional repercussions.

[5] There is no need for the regulator to impose safety standards on the cross-level security, since is optimal (cf. Lemma 1) for each target k to choose a level q_k that discourages indirect attacks in equilibrium.

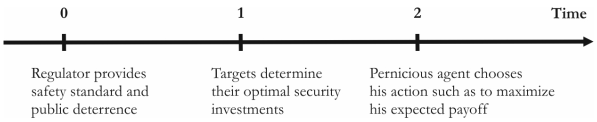

Figure 3: Timeline.

Targets. At time $t = 1$, the risk-neutral targets – anticipating a possible attack by the pernicious agent – decide about their respective investments in security. Target k's beliefs about the distribution of any *other* target's expected losses, $\tilde{L}_j$, in the event of a successful attack are independent and identically distributed (i.i.d.) according to the probability distribution F with support $\mathbb{R}_+$. Each target $k \in \mathcal{K}$ trades off its privately known expected loss L_k against the cost of the detection-capability tuple (p_k, q_k). This cost is described by the function $c : [0,1]^2 \to \mathbb{R}_+$, which is assumed to be increasing (with nonzero gradient), twice continuously differentiable, convex, and such that $c(0,0) = 0$, and the Inada conditions,

$$c(p,1) = c(1,q) = \infty, \quad \left.\frac{\partial c}{\partial p}\right|_{(0,q)} = \left.\frac{\partial c}{\partial q}\right|_{(p,0)} = 0,$$

are satisfied for all $p, q \in [0,1]$. The latter ensure the interiority of solutions: a small improvement in detection from a zero level is almost costless while the expense to ensure 100% detection is unbounded. All targets take their investment decisions simultaneously. They are effectively taking part in an all-pay contest with private information, the "loser" of which will be attacked by the pernicious agent, provided that it is individually rational for the latter to do so. The all-pay contest is nonstandard in the sense that the targets' actions are multidimensional, and that the "winner" is determined not based on the targets' "bids," i.e., their security investments, but based on the agent's own payoff assessment. Note also that every target tries its best to avoid being singled out for an attack. The expected payoffs from an attack on target k both from the agent's and the target's point of view are summarized in Table 1.

Regulator. At time $t = 0$, the regulator determines the safety standards in terms of the minimum (direct) detection probability $\rho \in [0,1)$ which is mandatory for all targets. A safety standard for the minimum indirect security level is not necessary, as will become clear later. The regulator also decides about an appropriate deterrent $\delta \geq 0$ in the form of a disutility to the pernicious agent in the event of a failed attack. Given a symmetric equilibrium strategy profile such that $p_k = p^*(L_k; \delta, \rho)$ for all $k \in \mathcal{K}$, the overall objective for the regulator is to maximize social welfare,[6]

$$W(\delta, \rho) = K \int_{\mathbb{R}_+} \Pi^*(\ell; \delta, \rho)\, dF(\ell) - D(\delta), \qquad (1)$$

where $\Pi^*(L; \delta, \rho)$ is the equilibrium payoff of a target with expected loss L. The increasing, convex, and continuously differentiable function $D : \mathbb{R}_+ \to \mathbb{R}_+$, with $D(0) = 0$, describes the cost to society of providing a deterrent. We discuss the solution to the regulator's problem in Section 4.

m	*Target Payoff:* $\pi_k^m(p_k, q_k)$	*Agent Payoff:* $\varphi_k^m(p_k, q_k)$
d	$-(1-p_k)L_k - c(p_k, q_k)$	$L_k - p_k(L_k + \delta)$
i	$-(1-\rho)(1-q_k)L_k - c(p_k, q_k)$	$L_k - (q_k + (1-q_k)\rho)(L_k + \delta)$
$\emptyset$	$-c(p_k, q_k)$	0

Table 1: Expected Payoffs Resulting from an Attack $a = (k, m)$. (NB: $\pi_j^m = 0$, $j \neq k$.)

3 Optimal Security Investment

To determine a perfect Bayesian equilibrium of the three-period game we proceed by backward induction, starting at time $t = 2$.

The Agent's Attack Policy ($t = 2$). Taking into account the possibility that the harmful device is detected in a security check spoiling an attack, the agent's expected payoffs φ_k^m given an attack on target k using attack mode m are given in Table 1. At $t = 2$, the agent chooses an attack policy $a^* = (k^*, m^*) \in \mathcal{K} \times \{\emptyset, d, i\}$ such that

$$a^* \in \arg\max_{(k,m)} \{\varphi_k^m\}, \qquad (2)$$

where φ_k^m are given in Table 1. Comparing the agent's motivation between attacking target k directly and indirectly we have that

$$\varphi_k^d \geq \varphi_k^i \quad \Leftrightarrow \quad p_k \leq q_k + (1-q_k)\rho, \qquad (3)$$

independent of both the expected loss L_k and the regulator's deterrent δ.

Security Investments ($t = 1$). After privately observing their respective expected losses, the targets noncooperatively determine their security investments, anticipating the agent's attack policy. Thus, given L_k target k seeks to determine $p_k \in [\rho, 1]$ and $q_k \in [0,1]$ so as to maximize its expected payoffs, given the other targets' strategies. For simplicity, we restrict attention to symmetric Bayes-Nash equilibria in which $p_k = p(L_k)$ and $q_k = q(L_k)$ for all $k \in \mathcal{K}$.

Lemma 1. *In a symmetric equilibrium the pernicious agent is indifferent between attacking a target directly or indirectly, i.e., $\varphi_k^d = \varphi_k^i$ for all $k \in \mathcal{K}$.*

[6] The pernicious agent's payoff is not included in the regulator's understanding of social welfare.

The intuition for the last result is that any excess investment in preventing a direct attack compared to preventing an indirect attack is wasted, as the agent always chooses the most promising attack mode. Lemma 1 therefore lays the foundation for determining the targets' equilibrium policies at $t = 1$, establishing an affine relation between direct and indirect security capabilities p_k and q_k. Indeed, relation (3) together with Lemma 1 implies that

$$q_k = \frac{p_k - \rho}{1-\rho} \quad (\leq p_k), \tag{4}$$

for all $k \in \mathcal{K}$, in equilibrium. Because of the fixed proportions in direct vs. indirect security investment, we can set

$$C(p) = c(p, \frac{p-\rho}{1-\rho}), \tag{5}$$

for any $p, \rho \in [0, 1)$. We are now ready to determine the targets' unique symmetric equilibrium strategies. For this, we first note that any target can always guarantee itself an attack-free existence by investing in the "full security" level $\bar{p}(L) = L/(L+\delta)$ which ensures that the agent's attack payoff vanishes, as $\varphi = L - \bar{p}(L)(L+\delta) = 0$. Using the convention that $\sup \emptyset = 0$, let

$$r(L) = \sup\left\{\hat{r} \in [0,1] : C\left(\tfrac{L}{L+\delta}\right) - (1-\hat{r})L \geq C(\hat{r})\right\}$$

denote a "critical security" level that renders a loss-L target indifferent between no attack at $\bar{p}(L)$ and an attack at $r(L)$. It is clear that $r(L)$ never exceeds $\bar{p}(L)$, and that it can only be positive if L lies strictly between zero and the expected loss $\hat{L}$ above which investment in full security becomes prohibitively expensive, determined by $\hat{L} = C(\bar{p}(\hat{L}))$.[7] Furthermore, when $r(L)$ vanishes for some $L > 0$, then for that loss it is not optimal to invest in full security.

3.1 First-Best Security Investment

Consider the socially optimal policy at time $t = 1$. If the regulator or social planner has full control over each target's security investment, then the resulting optimal security-investment policy is termed "first best" and denoted by $(p^{\text{FB}}(L), q^{\text{FB}}(L))$, $L \geq 0$. It is subject to the security infrastructure (δ, ρ) as determined in the preceding period (i.e., at $t = 0$; cf. Section 4.1 below). By Lemma 1 it is

$$q^{\text{FB}}(L) \equiv \frac{p^{\text{FB}}(L) - \rho}{1-\rho}, \tag{6}$$

so that we can restrict attention to the determination of $p^{\text{FB}}(L)$. In equilibrium, because of the convexity of the cost $C(p)$ for security p, the agent will be attacking the target with the highest expected loss that is available. Otherwise, the planner could save by decreasing the security of a high-loss target without increasing the likelihood of an attack, which contradicts the optimality of a policy $p(L)$ that would lead to a nonmonotonic agent payoff $\varphi(L) = L - p(L)(L+\delta)$. Given (δ, ρ), the social planner's problem is therefore [8]

$$\max_{p(\cdot)\in[0,1]} E\left[-G(\tilde{L})\left(1-p(\tilde{L})\right)\tilde{L} - C(p(\tilde{L}))\,\middle|\,\delta,\rho\right]$$
$$\text{s.t.} \quad p(L) \geq \rho, \quad \dot{\varphi}(L) \geq 0, \quad L \geq 0. \tag{7}$$

This variational problem can be solved explicitly using optimal control theory [13, 19]. Its solution is very simple when the monotonicity constraint is not binding, in which case the marginal cost of increasing the security level further is equal to the expected loss L times the probability $G(L)$ of this loss being the highest (securing the agent's attack interest).

Proposition 1 (First-Best Security Investment). *The unique first-best security-investment policy $(p^{FB}(L), q^{FB}(L))$, $L \geq 0$, is such that $p^{FB}(L) = \rho$ for all $L \in [0, L_\rho]$, and*

$$p^{FB}(L) = \begin{cases} C'^{-1}(G(L)L), & 1 > [p^{FB}(L)(L+\delta)]', \\ 1 - \dot{p}^{FB}(L)(L+\delta), & \textit{otherwise}, \end{cases} \tag{8}$$

for all $L \geq 0$, where $0 < L_\rho \equiv \sup\{\ell \geq 0 : G(\ell)\ell \leq C'(\rho)\} \leq \bar{L}$, and $q^{FB}(L)$ is obtained from Eq. (6).

The first solution in Eq. (8) determines the optimal security investment based on the marginal cost, whereas the second is such that the agent is indifferent between attacking targets of varying losses. The first policy is implemented as long as the agent's payoff is increasing, otherwise the second policy is used. The second policy, which is, for notational convenience, determined only implicitly, is such that

$$p^{\text{FB}}(L) = 1 - (1 - p^{\text{FB}}(\hat{L}))\frac{\hat{L}+\delta}{L+\delta}$$

provided that $\varphi^{\text{FB}}(L)$ is never increasing in some right-neighborhood of the point $\hat{L} \geq 0$ and the value $p^{\text{FB}}(\hat{L})$ is known.

Corollary 1. *The first-best security-investment policy $p^{FB}(L)$ is nondecreasing in ρ and generally nonmonotonic in δ, whereas $q^{FB}(L)$ is generally nonmonotonic in ρ and δ.*

[7] By l'Hôpital's rule the definition of $r(L)$ implies that $\dot{r}(0) = 1$, so that $r(L) > 0$ in a right-neighborhood of $L = 0$.

[8] Despite the apparent nonoptimality of a policy $p(L)$ does not lead to a nondecreasing agent payoff the monotonicity of $\varphi(L)$ needs to be imposed as a constraint in (7) to ensure that $G(L) = F^{K-1}(L)$ remains the distribution that correctly reflects the agent's attack policy.

Proposition 2 (**First-Best vs. Second-Best).** *Let $L_0 = \delta\rho/(1-\rho) < \bar{L}$. In equilibrium, any target with a loss $L > L_0$ underinvests in security compared to the first-best solution.*

3.2 Security-Investment Equilibrium

The following result summarizes the agent's and the firms' equilibrium behavior as a function of their respective expected losses.

Proposition 3 (**Security-Investment Equilibrium).** *Let $L_1 \geq L_0 = \delta\rho/(1-\rho)$ and $\omega = (L_0, L_1)$.*

(i) If $L_0 > \bar{L}$, then the pernicious agent does not attack any target, and the unique (symmetric) security-investment equilibrium is $(p(L), q(L)) \equiv (\rho, 0)$.

(ii) If $L_0 \leq \bar{L}$, then the unique (symmetric) security-investment equilibrium $(p(L), q(L))$, with $q(L) = \frac{p(L)-\rho}{1-\rho}$, solves[9]

$$\frac{\dot{p}(L)}{1-p(L)} + \frac{\hat{g}(L;\omega)L}{C'(p(L)) - \hat{G}(L;\omega)L - \lambda(L;\omega)} = \frac{1}{L+\delta}, \quad (9)$$

with $p(L_0) = \rho$, $p(L_1) = \bar{p}(L_1)$, where $\lambda(L;\omega) = \left[C'(\rho) - L\,\frac{d}{dL}(G(L;\omega)(L+\delta))\right]_+$.

The function $\varphi(L) = L - p(L)(L+\delta)$ is the agent's equilibrium payoff from attacking an agent with expected loss L. The agent is willing to attack if and only if this payoff is nonnegative.

Corollary 2. *The agent's equilibrium payoff $\varphi(L)$ is nondecreasing, and therefore*

$$m \neq \emptyset \quad \Leftrightarrow \quad \varphi(L) \geq 0 \quad \Leftrightarrow \quad L \geq L_0, \qquad (10)$$

i.e., the agent attacks only targets with losses of at least L_0.

The loss threshold $L_0 = \delta\rho/(1-\rho)$ is entirely determined by the regulatory policy (δ, ρ). It is interesting to note that the two instruments are extensive complements at the origin, in the sense that it is impossible to deter the agent from attacking a positive-loss target using only one of the two instruments (except when setting $\rho = 1$). By setting $L_0 = \bar{L}$ the regulator can ensure that no attacks take place in equilibrium, but the social cost of implementing such a "no-risk" policy may be prohibitive.

Corollary 3. *The targets' equilibrium strategies $p(L)$ and $q(L)$ are generally nonmonotonic in L.*

The nonmonotonicity of the targets' equilibrium strategies means that it is possible that higher-risk targets invest less in security than lower-risk targets. More specifically, it is possible to show that the strategies are always nondecreasing for losses in the interval $[0, L_0 + \varepsilon]$, where ε is some positive constant. Yet, as the following simple example illustrates, it is entirely possible that the highest-risk targets also have the lowest security. The reason is that the negative-externality exerted by the security efforts of lower-risk targets makes an attack on the higher-risk targets so likely that spending additional funds on security measures becomes undesirable.

Example 1. Let the public deterrent $\delta \geq 0$ and the safety standard $\rho \in [0, 1)$ be given. We assume that the cost of security is linear, so that $c(p,q) = \alpha p + \beta q$, where α, β are nonnegative constants with $\alpha + \beta > 0$. By Lemma 1 and Eqs. (4)–(5) we can restrict attention to direct security investments with the cost function $C(p) = c(p, \frac{p-\rho}{1-\rho}) = \alpha p + \beta\frac{p-\rho}{1-\rho} \equiv \gamma p - \frac{\beta\rho}{1-\rho}$, where $\gamma = \alpha + \frac{\beta}{1-\rho} > 0$. With this, the initial value problem (9) yields the agent's expected payoff from attacking a target of loss L,

$$\varphi(L) = \left(\exp\left[\int_{L_0}^{L} \frac{g(\ell)\ell\, d\ell}{\gamma - G(\ell)\ell}\right] - 1\right)\delta,$$

for all $L \in [0,1]$, given $L_0 = \delta\rho/(1-\rho)$. In the case of a uniform loss distribution, with $G(L) \equiv F^{K-1}(L) \equiv L^{K-1}$ and $g(L) = \dot{G}(L) = (K-1)L^{K-2}$, we therefore obtain

$$\varphi(L) = \left(\left(\frac{\gamma - L_0^K}{\gamma - L^K}\right)^{(K-1)/K} - 1\right)\delta,$$

and thus,

$$\begin{aligned} p(L) &= \max\left\{\rho, 1 - \frac{\delta}{L+\delta}\left(\frac{\gamma - L_0^K}{\gamma - L^K}\right)^{(K-1)/K}\right\}, \\ q(L) &= \frac{p(L) - \rho}{1-\rho}, \end{aligned}$$

for all $L \in [0,1]$, given the interesting case where $L_0 = \delta\rho/(1-\rho) \leq \bar{L} = 1$. It is straightforward to show that $p(L)$ is nondecreasing on $[0,1]$ if and only if $\gamma \geq 1 + (1+\delta)K$. Thus, if the marginal cost of security, γ, is relatively small or the number of targets, K, is large (so that the previous inequality is not satisfied), then the equilibrium safety $p(L)$ will decrease in L in the high-loss region. In other words, some targets with intermediate losses like to invest more in safety than high-loss targets. □

[9]A dot denotes total derivatives with respect to the independent variable L, i.e., $\dot{\varphi}(L) \equiv d\varphi(L)/dL$.

Proposition 4 (**Comparative Statics).** *Let $(p(L), q(L))$, $L \geq 0$, be a symmetric security-investment equilibrium as in Proposition 3, and $\varphi(L)$ be the agent's equilibrium payoff conditional on attacking a target of loss L.*

(i) $p(L)$ is nonincreasing in K, nondecreasing in ρ, and generally nonmonotonic in δ.

(ii) $q(L)$ is nonincreasing in K, and generally nonmonotonic in ρ, δ.

(iii) $\varphi(L)$ is nondecreasing in K, nonincreasing in ρ, and nondecreasing in δ for small $\delta > 0$ (but generally nonmonotonic for larger δ's).

Part (iii) implies that a slight increase of a small deterrent may have the perverse effect to increase the agent's expected payoffs and thus elevate the likelihood of an attack. This somewhat counter-intuitive result stems from the fact that it decreases the targets' incentives to invest in security. The result is actually opposite for small L.

In addition, we find that increasing the deterrent δ increases the attacker's payoff for small losses. Moreover, increasing the safety standards decreases the attacker's payoff for all loss levels and increases the investment in safety. Finally, if $\rho \geq \bar{L}/(\bar{L} + \delta)$, then there will be no attack.

Overall, the deterrent is bad for targets with losses $L \geq L_0$. A further increase is immaterial for targets with losses $L < L_0$, since they will not be attacked at all. Their safety is set at the minimum standard ρ. Both instruments have the effect of reducing the number of targets. The benefits for society are concentrated in the margin. The deterrent is bad for high risks, the safety standard bad for low risks.

Corollary 4. *As the number of targets K goes to infinity, the targets' security investments become minimal, as $\lim_{K\to\infty}(p(L), q(L)) = (\rho, 0)$.*

The incentives to acquire security diminish as there are more and more other targets.

4 Safety Standards and Public Deterrence

4.1 Efficient Regulatory Policy

We now provide criteria for both the optimal safety standard and the most effective deterrent from a public-policy viewpoint.

Proposition 5 (**Optimal Regulation).** *The socially optimal safety standard $\hat{\rho}$ and deterrent $\hat{\delta}$ satisfy*

$$\frac{\hat{\delta}\hat{\rho}\, F^{K-2}(L_0) f(L_0)}{(1-\hat{\rho})^2} = C'(\hat{\rho}) + \frac{\int_{L_0}^{\bar{L}} \lambda(\ell;\hat{\delta},\hat{\rho}) dF(\ell)}{F(L_0)} \quad (11)$$

and

$$\frac{\hat{\delta}\hat{\rho}^2 F^{K-1}(L_0) f(L_0)}{1-\hat{\rho}} = \frac{D'(\hat{\delta})}{K} + \int_{L_0}^{\bar{L}} \frac{G(\ell)\ell + \lambda(\ell;\hat{\delta},\hat{\rho})}{\ell+\hat{\delta}}\, p^*(\ell;\hat{\delta},\hat{\rho}) dF(\ell), \quad (12)$$

where the Lagrange multiplier $\lambda(L;\hat{\delta},\hat{\rho})$ is given by[10]

$$\lambda(L;\hat{\delta},\hat{\rho}) = \Big[C'(p^*(L;\hat{\delta},\hat{\rho})) - G(L)L - (L+\hat{\delta})g(L)L\Big]_+ \geq 0, \quad (13)$$

for all $L \in [0, \bar{L}]$, and $p^*(L;\hat{\delta},\hat{\rho})$ is part of the targets' symmetric security-investment equilibrium in Prop. 3.

For any given $\rho > 0$ there is a unique solution, since the left-hand side is increasing and the right-hand side is decreasing. Moreover, the right-hand side is bounded whereas the left-hand side varies over the entire space $\mathbb{R}_+$.

The effect of increasing public deterrence δ on symmetric equilibrium investment is generally ambiguous. A higher level of deterrence may both increase and decrease equilibrium security investment of different targets, depending on their private expected losses. To show this consider the following continuation of our earlier example.

Example 2. Assume first $2 + \delta \leq \alpha$, so that $\lambda(L) = 0$ on $[0, 1]$. Then $L_0^{K-1} = (1-\rho)\alpha$ and

$$\rho L_0^K = d\delta + \int_{L_0}^1 \frac{\ell^K}{\ell+\delta}\left(1 - \frac{\delta}{\ell+\delta}\left(\frac{\alpha - L_0^K}{\alpha - \ell^K}\right)^{(K-1)/K}\right) d\ell$$

characterize the socially optimal choice of δ and ρ.

There is also the idea of discouraging underinvestment through safety standards. More specifically, one can allow for penalties to the targets in the case of a successful attack, reinternalizing the social externality. Such punishments EW only effective if the targets remain liquid after a successful attack and do not seem to correspond to common practice.

4.2 Other Policy Choices

To prevent socially inefficient overinvestment in security, the social planner could cap the targets' losses by providing a guarantee of last resort. Losses are also bounded from above by the targets' limited liability.

In the case of inefficient underinvestment the social planner could impose fines in the case of a loss to make

[10]For any $x \in \mathbb{R}$ we define by $[x]_+ = \max\{0, x\}$ the nonnegative part of x.

the targets internalize the harmful externalities to society. However, such punishments may not be effective or desirable due to the targets' limited liability and other reasons (in the public eye the government should help targets in a misfortune rather than punish the targets on top of it).

5 Conclusion

In this paper we have examined a SIDS problem with first and second-order payoff interdependencies. The equilibrium security investments depend on the costs of achieving a high detection capability as well as on the public policies implemented by the social planner. If a high-loss target finds it very expensive to install security systems it will balance expected losses against its investment in providing safety while on the other hand accepting the fact that it likely still represents a most attractive target to the terrorist who will be compelled to attack. If on the other hand it is sufficiently cheap for the high-loss target to install security systems, it is in its best interest to provide "full" security in the sense that the terrorist most likely responds by not attacking the target.[11] A target's decision is influenced by its expectation of the *lowest* direct security level that is implemented in equilibrium as this determines the likelihood of an indirect attack. We show that in equilibrium the pernicious agent cannot have a strict preference for such indirect attacks. Nevertheless, the externalities implied by the targets' interconnectedness influence the equilibrium significantly: the lower the expected lowest direct security level, the more likely a target will opt for full protection. In fact, in our considerations we restrict attention to the situation in which there are (at least weak) economies of scale (and complementarities) in the security technologies, which results in threshold policies switching between full and minimum security investment. There is some empirical evidence for economies of scale in the provision of security, at least over a certain range [21]. We find that interestingly the resulting equilibrium investment policies do *not* need to be monotonic in the expected private losses. There can be many switching points as these private losses increase, reflecting the delicate balance of benefits when comparing the two options.

In discussing policy options one has to examine both the private and social welfare aspects of investing in security. If from a social welfare perspective each agent overinvests in protection in a competitive environment because each one fears being a target of attack, then some public involvement is necessary. This can take the form of a well-enforced regulation or standard or economic incentives such as a subsidy or fine. — For once, a social planner can provide deterrence in the form of commitment to punishment using legal means or retaliation. Such public deterrence may have an ambiguous effect on private security investments. On the one hand, it may decrease target spending since public deterrence in effect constitutes an imperfect substitute for private security systems; on the other hand, it may also increase target spending, since it tends to narrow the 'cost gap' for achieving full security. Safety standards are a relatively weak instrument for the public policy maker, since it may lead to inefficient overinvestment in security, especially for the provision of private intertarget (i.e., indirect) security. We find that safety standards tend to increase private investments even beyond the prescribed level: it encourages full security investment at least for those close to a previous indifference point. Punishments conditional on an attack being carried out successfully may work in theory, if the punishment can be made contingent on the loss. But it is difficult to enforce contingencies on the social loss portion. In addition, the target is subject to limited liability. Note that the public provision of deterrence compensates for inefficiencies arising from asymmetries in the targets' loss distributions as well as differences in their cost of installing security systems.

To provide incentives for security investment in a situation where firms face security interdependencies is a complex task.[12] The two main instruments, namely safety standards and public deterrent against the intruder, require a credible commitment by the planner. Regulatory commitment is necessary for the firms to agree to the irreversible security investments. The optimal level of commitment, which could be achieved by the regulator by self-imposing penalties for the ex-post deviation from announced standards and deterrence guarantees, may be subject to optimization itself (cf. [20]).

[11] Under complete information targets may enter into a direct all-pay contest with the low-loss target leading generally to a mixed-strategy equilibrium unless one of the targets drops out of the contest by selecting one of its outside options. When entering a rent-seeking contest (under complete information) at most one target will do so at a gain [2]. The other target will, in equilibrium, be indifferent between the expected payoffs it receives from the contest and the expected payoffs from its best outside option. The presence of private information (such as the targets' respective expected losses) tends to restore a pure-strategy (Bayes-)Nash equilibrium [8].

[12] The main focus here was on physical security with interdependencies. There is a link to the literature of software security [17].

References

[1] AUMANN, R.J. (1976) "Agreeing to Disagree," *Annals of Statistics*, Vol. 4, No. 6, pp. 1236–1239.

[2] BAYE, M.R., KOVENOCK, D., DE VRIES, C.G. (1996) "The All-Pay Auction with Complete Information," *Economic Theory*, Vol. 8, No. 2, pp. 291–305.

[3] CAULCY, J., IM, E.I. (1988) "Intervention Policy Analysis of Skyjackings and Other Terrorist Incidents," *American Economic Review*, Vol. 78, No. 2, pp. 27–31.

[4] CESARI, L. (1983) *Optimization – Theory and Applications: Problems with Ordinary Differential Equations*, Springer, New York, NY.

[5] ENDERS, W., SANDLER, T. (1995) "Terrorism: Theory and Applications," in: *Handbook of Defense Economics*, Vol. 1, K. Hartley and T. Sandler (eds.), Elsevier Science, Amsterdam, NL, pp. 213–249.

[6] HEAL, G., KUNREUTHER, H. (2003) "You Only Die Once: Managing Discrete Interdependent Risks," NBER Working Paper 9885, National Bureau of Economic Research, Cambridge, MA.

[7] HIRSHLEIFER, J. (1988) "The Analytics of Continuing Conflict," *Synthese*, Vol. 76, No. 2, pp. 201–233.

[8] KAPLAN, T., LUSKI, I., SELA, A., WETTSTEIN, D. (2002) "All-Pay Auctions with Variable Rewards," *Journal of Industrial Economics*, Vol. 50, No. 4, pp. 417–430.

[9] KEOHANE, N.O., ZECKHAUSER, R. (2003) "The Ecology of Terror Defense," *Journal of Risk and Uncertainty*, Vol. 26, No. 2/3, pp. 201-229.

[10] KUNREUTHER, H., HEAL, G. (2002) "Interdependent Security: The Case of Identical Agents," *Journal of Risk and Uncertainty*, Vol. 26, No. 2/3, pp. 231–249.

[11] LAPAN, H.E., SANDLER, T. (1988) "To Bargain or Not To Bargain: That is the Question," *American Economic Review*, Vol. 78, No. 2, pp. 16–21.

[12] PATÉ-CORNELL, E., GUIKEMA, S. (2003) "Probabilistic Modeling of Terrorist Threats: Systems Analysis Approach to Setting Priorities Among Countermeasures," *Military Operations Research*, Vol. 7, No. 4, pp. 5–20.

[13] PONTRYAGIN, L.S., BOLTYANSKII, V.G., GAMKRELIDZE, R.V., MISHCHENKO, E.F. (1962) *The Mathematical Theory of Optimal Processes*, Wiley Interscience, New York, NY.

[14] SANDLER, T., LAPAN, H.E. (1988) "The Calculus of Dissent: An Analysis of Terrorists' Choice of Targets," *Synthese*, Vol. 76, No. 2, pp. 245–261.

[15] SELTEN, R. (1977) "A Simple Game Model of Kidnapping," in: *Mathematical Economics and Game Theory,* R. Henn and O. Moeschlin (eds.), Lecture Notes in Economics and Mathematical Systems, Vol. 141, Springer-Verlag: Berlin, pp. 139-156.

[16] SMITH, V.L. (1968) "Optimal Insurance Coverage," *Journal of Political Economy*, Vol. 76, No. 1, pp. 68–77.

[17] TUNCA, T.I. (2006) "Network Software Security and User Incentives," *Management Science*, Vol. 52, No. 11, pp. 1703–1720.

[18] WALTER, W. (1998) *Ordinary Differential Equations*, Springer, New York, NY.

[19] WEBER, T.A. *Optimal Control Theory with Applications in Economics*, MIT Press, Cambridge, MA.

[20] WEBER, T.A. (2014) "A Continuum of Commitment," *Economics Letters*, Vol. 124, No. 1, pp. 67–73.

[21] WEIN, L.M. (2008) "Homeland Security: From Mathematical Models to Implementation: The 2008 Philip McCord Morse Lecture," *Operations Research*, Vol. 57, No. 4, pp. 801–811.

Appendix: Proofs

Proof of Lemma 1. Let $(p(L), q(L))$, $L \in [0, \bar{L}]$, be a symmetric Bayes-Nash equilibrium describing the targets' security investments at time $t = 1$. The pernicious agent strictly prefers attacking target k directly to attacking it indirectly if and only if $\varphi_k^d(L_k) > \varphi_k^i$, which by (3) is equivalent to

$$p_k < q_k + (1 - q_k)\rho.$$

Thus, target k could save costs by decreasing q_k somewhat, since it does not increase the probability of being attacked successfully. Some of the cost savings can then be reinvested to augment the direct security level p_k by some finite amount. Hence, target k can profitably deviate from the equilibrium strategy profile. The proof that there is also a deviation when the agent strictly prefers a direct attack proceeds in a similar manner, so that in equilibrium an attacking agent must (at least *ex ante*) be indifferent over whether to attack target k directly or indirectly, i.e., $\varphi_k^d = \varphi_k^i$ for all $k \in \mathcal{K}$. ■

Proof of Proposition 1. We first rewrite the planner's optimization problem with the agent's payoff as choice variable, similar to our proof of Proposition 3,

$$\begin{aligned} &\max_{\varphi(\cdot)\geq 0} \int_0^{\bar{L}} \left(-\hat{G}(\ell)\frac{\varphi(\ell)+\delta}{\ell+\delta} - \hat{C}(\varphi(\ell),\ell) \right) dF(\ell) \\ &\text{s.t.} \quad \frac{L-\varphi(L)}{L+\delta} \geq \rho, \quad \dot{\varphi}(L) \geq 0, \quad L \in [0, \bar{L}], \end{aligned} \tag{14}$$

where $\hat{G}(L) \equiv G(L)L$. By introducing the control $u = \dot{\varphi}$, it is possible to formulate it equivalently as a standard optimal control problem,

$$\begin{aligned} &J(u) = \int_0^{\bar{L}} \left(-\hat{G}(\ell)\frac{\varphi(\ell)+\delta}{\ell+\delta} - \hat{C}(\varphi(\ell),\ell) \right) f(\ell)d\ell \longrightarrow \max_{u(\cdot)} \\ &\text{s.t.} \quad \dot{\varphi}(L) = u(L), \quad \varphi(0) = 0, \quad u(L) \geq 0, \\ &\qquad \frac{L-\varphi(L)}{L+\delta} \geq \rho, \quad L \in [0, \bar{L}]. \end{aligned} \tag{15}$$

The corresponding Hamilton-Pontryagin function is

$$H(L, \varphi, u, \psi) = \psi u - \left(\hat{G}(L)\frac{\varphi+\delta}{L+\delta} + \hat{C}(\varphi, L) \right) f(L),$$

where $\psi = \psi(L)$ is an adjoint variable. We first consider a relaxed version of this problem, without the state constraint $(L - \varphi(L))/(L + \delta) \geq \rho$. A solution to this relaxed problem exists [4], and can be obtained using the Pontryagin maximum principle [13]. Indeed, given an optimal state-control trajectory $(\varphi^{\text{FB}}(L), u^{\text{FB}}(L))$, $L \in [0, \bar{L}]$, the Pontryagin maximum principle implies that there exists an absolutely continuous function $\psi(L)$ such that the following conditions are satisfied.

1. Adjoint equation:

$$\begin{aligned} \dot{\psi}(L) &= -\left.\frac{\partial H}{\partial \varphi}\right|_{(L,\varphi^{\text{FB}}(L),u^{\text{FB}}(L),\psi(L))} \\ &= \left(\hat{G}(L) - C'\left(\frac{L-\varphi^{\text{FB}}(L)}{L+\delta} \right) \right) \frac{f(L)}{L+\delta}, \end{aligned}$$

   for all $L \in [0, \bar{L}]$.

2. Transversality condition:

$$\psi(\bar{L}) = 0.$$

3. Maximality condition:

$$u^{\text{FB}}(L) \in \arg\max_{u\geq 0} H(L, \varphi^{\text{FB}}(L), u, \psi(L)),$$

   for almost all $L \in [0, \bar{L}]$.

From the maximality condition we obtain that $u^{\text{FB}}(L)\psi(L) = 0$ a.e. on $[0, \bar{L}]$. Thus, if $u^{\text{FB}}(L) > 0$ a.e. on some interval, then $\dot{\psi}(L) = 0$ on the same interval, so that the adjoint equation implies

$$p^{\text{FB}}(L) \equiv \frac{L-\varphi^{\text{FB}}(L)}{L+\delta} = C'^{-1}(\hat{G}(L)), \tag{16}$$

which is a nondecreasing function.[13] If $u^{\text{FB}}(L) = 0$ a.e. on some interval $[\hat{L}, \hat{L} + \varepsilon] \subset [0, \bar{L}]$ (for some $\varepsilon > 0$), then

$$\begin{aligned} \dot{\varphi}(L) &= \frac{d}{dL}\left(L - p^{\text{FB}}(L)(L+\delta) \right) \\ &= 1 - p^{\text{FB}}(L) - \dot{p}^{\text{FB}}(L)(L+\delta) = 0, \end{aligned}$$

which is a (separable) ordinary differential equation, with the following unique increasing solution (taking into account the known initial value $p^{\text{FB}}(\hat{L})$):

$$p^{\text{FB}}(L) = 1 - \left(1 - p^{\text{FB}}(\hat{L}) \right) \frac{\hat{L}+\delta}{L+\delta} \tag{17}$$

for all $L \in [\hat{L}, \hat{L} + \varepsilon]$. We therefore conclude that the relaxed version of problem (15) has in fact the same solution as the original problem (15), as long as the relaxed solution $p^{\text{FB}}(L)$ is at least ρ, or, equivalently, as long as $L \geq L_\rho$, where

$$L_\rho = \sup\{\ell \in [0, \bar{L}] : p^{\text{FB}}(\ell) \leq \rho\},$$

given the convention that $\sup \emptyset = 0$. Note also that $\hat{G}(0) = 0 < \hat{C}'(0)$, so that $p^{\text{FB}}(L) = \rho$ in a right-neighborhood of $L = 0$ and thus $L_\rho > 0$. Moreover, on $[0, L_\rho]$ we have that $u^{\text{FB}}(L) = \dot{\varphi}(L) = 1 - p^{\text{FB}}(L) > 0$ (for the relaxed problem), so that $\psi(L) = 0$. By continuity of ψ we therefore conclude that

$$\psi(L_\rho) = 0.$$

In addition, the threshold L_ρ can now be expressed in terms of the primitives of the problem,

$$L_\rho = \sup\{\ell \in [0, \bar{L}] : \hat{G}(\ell) \leq C'(\rho)\}.$$

Combining (16) and (17), together with the fact that the second solution simplifies because (independent of $\hat{L} > L_\rho$)

$$\left(1 - p^{\text{FB}}(\hat{L}) \right) \frac{\hat{L}+\delta}{L+\delta} = (1-\rho)\frac{L_\rho+\delta}{L+\delta},$$

we obtain the claims of the proposition. ■

Proof of Corollary 1. We first note that L_ρ as defined in Proposition 1 is nondecreasing in ρ by convexity of C and constant in δ. Yet, as can be seen in the proof of that proposition, L_ρ has no influence on the

[13] When C is not strictly convex, then it may not be possible to take the (multivalued) inverse of C', and the solution (17) is likely to be optimal.

solution of the relaxed solution, and thus the solution p^{FB} is weakly increasing in ρ where it is constant and otherwise essentially unaffected by ρ. The parameter δ has no effect on the first policy in (8). An increase in δ leads to an increase of the second policy at loss L if and only if $(L_\rho + \delta)/(L+\delta) > \delta$, leading to a generic nonmonotonicity in δ. The claims about $q^{\text{FB}}(L)$ now follow by considering Eq. (6). ■

Proof of Proposition 2. By Corollary 2 the agent's payoff $\varphi(L)$ is nondecreasing in a security-investment equilibrium $(p(L), q(L))$, $L \in [0, \bar{L}]$, and increasing for all $L > L_0$. For $L_k > L_0$, any target k's optimality condition for its payoff maximization problem in the proof of Proposition 3), the probability $p(L_k)$, which solves

$$\max_{p_k \in [\rho,1]} \left\{ -G(\varphi^{-1}(L_k - p_k(L_k+\delta)))(1-p_k)L_k - C(p_k) \right\},$$

can be written in the form

$$\frac{g(L_k)}{\dot{\varphi}(L_k)}(1-p(L_k))(L_k+\delta)L_k + G(L_k)L_k = C'(p(L_k)). \quad (18)$$

This condition states that the marginal cost of security for any agent k is in equilibrium equal to the benefit of that from the decreased probability of attack on target k and the increased probability of uncovering an attack. Let $\varphi^{\text{FB}}(L_k)$ be the agent's expected payoff of attacking target k when the first-best security policy $(p^{\text{FB}}, q^{\text{FB}})$ is implemented for all targets. Assuming that $\dot{\varphi}^{\text{FB}}(L_k) > 0$, the social planner's optimality condition (see Eq. (16) in the proof of Proposition 1) becomes

$$G(L_k)L_k = C'(p^{\text{FB}}(L_k)). \quad (19)$$

The marginal cost of the socially optimal security investment is equal to the decrease in loss due to the improved detection capability. Combining Eqs. (18) and (19) yields

$$\begin{aligned} C'(p^{\text{FB}}(L_k)) &= C'(p(L_k)) - \frac{g(L_k)}{\dot{\varphi}(L_k)}(1-p(L_k))(L_k+\delta)L_k \\ &> C'(p(L_k)), \end{aligned}$$

which, for the cost function C is by assumption convex, implies that

$$p^{\text{FB}}(L_k) < p(L_k). \quad (20)$$

Note now that for any (differentiable) security policy $p(L) < 1$, $L > L_0$, it is by definition $\varphi = L - p(L)(L+\delta)$, so that

$$\dot{p}(L) = \frac{1-p(L)}{L+\delta} - \dot{\varphi}(L) \geq \frac{1-p(L)}{L+\delta},$$

for all $L \in (L_0, \bar{L}]$, provided that $\dot{\varphi}(L) \geq 0$. This means that at a loss $L \in (L_0, \bar{L})$ the steepest increase of any security policy that does not decrease the agent's payoff has a slope of $(1 - p(L))/(L + \delta)$. Hence, when $\dot{\varphi}(L_k) = 0$, it is not possible for the second-best policy to ever catch up with the first-best policy on $[0, L_k]$, which (taking into account that under the second-best security-investment policy the agent's equilibrium payoff from attacking a target is actually increasing for losses exceeding L_0) implies that inequality (20) must hold for all $L_k \in (L_0, \bar{L}]$, completing our proof. ■

Proof of Proposition 3. (i) Note first that when $\rho = 1$, the unique security-investment equilibrium strategy is $(p(L), q(L)) \equiv (1, 0)$ and there is nothing to prove. We now assume that $(\delta, \rho) \in \mathbb{R}_+ \times [0, 1)$, and consider target 1's choice of (p_1, q_1) when all other targets play according to the symmetric strategy $(p(L), q(L))$. By Lemma 1 we obtain that $q(L) = (p(L) - \rho)/(1 - \rho)$ for all $L \in [0, \bar{L}]$ and $q_1 = (p_1 - \rho)/(1 - \rho)$. Let $\varphi_k = \varphi(L_k)$ be the agent's expected payoff from attacking target $k \in \{2, \ldots, K\}$ and let

$$\varphi_1 = L_1 - p_1(L_1 + \delta) \quad (21)$$

be his expected payoff from attacking target 1, which has an expected loss of $L_1 \in [0, \bar{L}]$. Note that if $\rho \geq \bar{L}/(\bar{L} + \delta)$ (or, equivalently, if $L_0 \leq \bar{L}$), then $p_1 \geq \rho$ implies that $\varphi_1 < 0$ for all $L_1 \in [0, \bar{L})$. Hence, using Eq. (4) we obtain $(p(L), q(L)) \equiv (\rho, 0)$. (ii) Using the notation established in the proof of part (i), target 1's expected payoff $\Pi(p_1, q_1)$ is equal to

$$-\mathbb{P}\left(\max\{\varphi(\tilde{L}_2), \ldots, \varphi(\tilde{L}_K)\} \leq \varphi_1\right)(1-p_1)\,L_1 - c(p_1, q_1).$$

Using Eqs. (5) and (21) we obtain that this payoff can be equivalently rewritten in the form

$$\begin{aligned} \hat{\Pi}(p_1, L_1) &= -\mathbb{P}(\max\{\varphi(\tilde{L}_2), \ldots, \varphi(\tilde{L}_K)\} \\ &\leq L_1 - p_1(L_1+\delta))\,(1-p_1)\,L_1 - C(p_1). \end{aligned}$$

We now assume (and verify *ex post*) that the agent's equilibrium payoff $\varphi(L)$ is increasing in L (at least as long as $\varphi(L) > 0$). Then $\varphi(L)$ is invertible, and

$$\begin{aligned} &\mathbb{P}\left(\max\{\varphi(\tilde{L}_2), \ldots, \varphi(\tilde{L}_K)\} \leq \varphi_1\right) \\ &= \mathbb{P}\left(\tilde{L}_2, \ldots, \tilde{L}_K \leq \varphi^{-1}(\varphi_1)\right) = F^{K-1}(\varphi^{-1}(\varphi_1)). \end{aligned}$$

Thus, after setting $G = F^{K-1}$ and taking into account the regulatory constraint that $p_1 \geq \rho$, target 1's payoff-maximization problem can be written in the form

$$\begin{aligned} &\max_{\varphi_1 \geq 0} \left\{ -G(\varphi^{-1}(L_1 - p_1(L_1+\delta)))(1-p_1)L_1 - C(p_1) \right\} \\ &\text{s.t.} \qquad p_1 \geq \rho. \end{aligned} \quad (22)$$

The first-order necessary optimality condition for the problem (22) is, using the inverse function theorem and the fact that $L_1 = \varphi^{-1}(\varphi_1)$,

$$\frac{g(L_1)L_1}{\dot{\varphi}(L_1)}(1-p_1)(L_1+\delta) + G(L_1)L_1 - C'(p_1) + \lambda(L_1) = 0,$$

where $\lambda(L_1)$ is the Lagrange multiplier associated with the inequality constraint $p_1 \geq \rho$. By setting $L_1 = L$ and realizing that $\dot{\varphi}(L) = 1 - p - \dot{p}(L)(L + \delta)$, we obtain

$$\left(\frac{1}{L+\delta} - \frac{\dot{p}(L)}{1-p(L)}\right)^{-1} g(L)L + G(L)L = C'(p(L)) - \lambda(L), \quad (23)$$

where $g(L) \equiv \dot{G}(L)$. The complementary slackness condition $\lambda(L)(p(L) - \rho) = 0$, together with the fact that $\dot{p}(L) = 0$ on any interval where $p(L) = \rho$, yields that

$$\begin{aligned} \lambda(L) &= \left[C'(\rho) - G(L)L - g(L)L(L+\delta)\right]_+ \\ &= \left[C'(\rho) - L\,\frac{d\,(G(L)(L+\delta))}{dL}\right]_+ \geq 0. \end{aligned} \quad (24)$$

Eqs. (23)–(24) with the initial condition $p(L_0) = \rho$ are equivalent to the initial value problem (9) for $L \geq L_0$. For $L < L_0$ we obtain $(p(L), q(L)) = (\rho, 0)$ as in part (i). The monotonicity of $\varphi(L)$ is established in the proof of Corollary 2.[14] ■

[14] A standard result from the theory of ordinary differential equations (see, e.g., [18]) is that since the differential equation in (9) has – when written in the form $\dot{\varphi}(L) = h(\varphi(L), L)$ – a right-hand side which is Lipschitz-continuous in φ and uniformly bounded (the latter holds, because we know that $\varphi(L)$ lies in $[0, L] \subset [0, \bar{L}]$ and is therefore uniformly bounded), there exists a unique solution $\varphi(L)$, $L \in [L_0, \bar{L}]$, to the initial value problem (9). Even if $L_0 = 0$, there is no singularity at $L = 0$, since by assumption $C'(0) > 0$.

Proof of Corollary 2. We now verify the monotonicity of the agent's equilibrium payoff $\varphi(L)$. Without loss of generality, let us consider the case where $L_0 = 0$. In that case, the denominator of the right-hand side of the differential equation in (9) is positive, since by assumption $C'(0) > 0$. The latter continues to hold for small positive L, so that the right-hand side stays nonnegative. By assumption the distribution G has the support $[0, \bar{L}]$, so that the corresponding density $g = \dot{G}$ is positive there on $[0, \bar{L}]$. Hence, $\dot{\varphi}$ is positive and therefore $\varphi(L)$ increasing on $[L_0, \bar{L}]$. Note that it is impossible that $G(L)L > C'(p(L))$ in equilibrium, because the slope $\dot{\varphi}$ must be bounded from above, for φ is uniformly bounded (cf. Footnote 14). Relation (10) obtains immediately from the monotonicity of $\varphi(L)$, together with the initial condition in (9). ■

Proof of Corollary 3. Example 1 serves as a proof of the generic nonmonotonicity of the security-investment equilibrium $(p(L), q(L))$, $L \in [0, \bar{L}]$. ■

Proof of Proposition 4. Consider first the effect of an increase of the number of targets, from K to $\hat{K} > K$. The differential equation in (9) can be rewritten in the form

$$\begin{aligned}\dot{\varphi}(L) &= \frac{(K-1)(\varphi(L)+\delta)F^{K-2}(L)f(L)L}{C'(p(L)) - F^{K-1}(L)L} \\ &= \frac{(K-1)(\varphi(L)+\delta)}{\frac{C'(p(L))}{F^{K-1}(L)} - L}\,\frac{f(L)L}{F(L)}.\end{aligned}$$

Consider now the point (L, φ), and compare the slopes of the trajectories through that point for K and $\hat{K}$. Indeed, the slope of the latter trajectory is greater if and only if

$$\frac{\hat{K}-1}{\frac{C'(p(L))}{F^{\hat{K}-1}(L)} - L} > \frac{K-1}{\frac{C'(p(L))}{F^{K-1}(L)} - L}.$$

The last inequality holds, because $F^{\hat{K}-1} \leq F^{K-1}$. Thus, at any point in the state space (L, φ) the agent's payoff $\hat{\varphi}(L)$ for $\hat{K}$ targets increases faster than the agent's payoff $\varphi(L)$ for K targets. But the initial condition $\varphi(L_0) = 0$ is independent of K, so that both trajectories start at the same point $(L_0, 0)$. This implies that $\hat{\varphi}(L) > \varphi(L)$ for all $L > L_0$. Lastly, because of part (i) of Proposition 3, it is $\hat{\varphi}(L) = \varphi(L)$ for $L < L_0$. Therefore the agent's equilibrium payoff is nondecreasing in the number of available targets. This immediately implies that $p(L)$ and $q(L)$ are both nonincreasing in K.

We now consider the effect of an increase in the safety standard ρ. The integral representation of the initial value problem (9) is

$$\varphi(L) = \int_{L_0}^{L} \frac{(\varphi(\ell)+\delta)g(\ell)\ell\, d\ell}{C'((\ell - \varphi(\ell))/(\ell+\delta)) - G(\ell)\ell}, \qquad (25)$$

for all $L \in [L_0, \bar{L}]$. By the Leibniz rule we obtain

$$\begin{aligned}\frac{\partial \varphi(L)}{\partial \rho} &= -\frac{(\varphi(L_0)+\delta)g(L_0)L_0}{C'(\frac{L_0 - \varphi(L_0)}{L_0+\delta}) - G(L_0)L_0}\,\frac{\partial L_0}{\partial \rho} \\ &\quad + \int_{L_0}^{L} \frac{\partial}{\partial \rho}\,\frac{(\varphi(\ell)+\delta)g(\ell)\ell\, d\ell}{C'(\frac{\ell-\varphi(\ell)}{\ell+\delta}) - G(\ell)\ell} \\ &= -\frac{\delta^2 g(L_0)L_0}{(1-\rho)^2(C'(\rho) - G(L_0)L_0)}\end{aligned}$$

$$+ \int_{L_0}^{L} \frac{\partial \varphi(\ell)}{\partial \rho}\,\frac{\left(\left[C'(\zeta)+\zeta C''(\zeta)\right]\Big|_{\zeta = \frac{\ell-\varphi(\ell)}{\ell+\delta}} - G(\ell)\ell\right)g(\ell)\ell\, d\ell}{\left(C'(\frac{\ell-\varphi(\ell)}{\ell+\delta}) - G(\ell)\ell\right)^2},$$

which implies that

$$\frac{\partial \varphi(L)}{\partial \rho} \leq -\frac{\delta^2 g(L_0)L_0}{(1-\rho)^2 C'(\rho)} = -\frac{\delta^3 \rho g(\delta/(1-\rho))}{(1-\rho)^3 C'(\rho)} < 0,$$

for all $L \in [L_0, \bar{L}]$. In other words, the agent's payoff is decreasing in ρ for $L > L_0$, and vanishes for $L \leq L_0$. This implies that $p(L)$ is nondecreasing in ρ. Differentiating Eq. (4) with respect to ρ yields

$$\frac{\partial q(L)}{\partial \rho} = \frac{(1-\rho)(\partial p(L)/\partial \rho) - (1 - p(L))}{(1-\rho)^2},$$

indicating that it is possible for $q(L)$ to decrease in ρ while $p(L)$ increases. Example 2 provides a concrete instance of this effect.

Lastly, let us examine an increase in the deterrent δ. Differentiating the integral representation (25) of the initial value problem (9) gives

$$\begin{aligned}\frac{\partial \varphi(L)}{\partial \delta} &= -\frac{(\varphi(L_0)+\delta)g(L_0)L_0}{C'(\frac{L_0-\varphi(L_0)}{L_0+\delta}) - G(L_0)L_0}\,\frac{\partial L_0}{\partial \delta} \\ &\quad + \int_{L_0}^{L} \frac{\partial}{\partial \delta}\,\frac{(\varphi(\ell)+\delta)g(\ell)\ell\, d\ell}{C'(\frac{\ell-\varphi(\ell)}{\ell+\delta}) - G(\ell)\ell} \\ &= -\frac{g(L_0)L_0^2}{C'(\rho) - G(L_0)L_0}\end{aligned}$$

$$+ \int_{L_0}^{L} \frac{\left(C'(\frac{\ell-\varphi(\ell)}{\ell+\delta}) + \frac{(\varphi(\ell)+\delta)(\ell-\varphi(\ell))}{(\ell+\delta)^2} C''(\frac{\ell-\varphi(\ell)}{\ell+\delta}) - G(\ell)\ell\right)g(\ell)\ell\, d\ell}{\left(C'(\frac{\ell-\varphi(\ell)}{\ell+\delta}) - G(\ell)\ell\right)^2}$$

$$+ \int_{L_0}^{L} \frac{\partial \varphi(\ell)}{\partial \delta}\,\frac{\left(C'(\frac{\ell-\varphi(\ell)}{\ell+\delta}) + \frac{\varphi(\ell)+\delta}{\ell+\delta} C''(\frac{\ell-\varphi(\ell)}{\ell+\delta}) - G(\ell)\ell\right)g(\ell)\ell\, d\ell}{\left(C'(\frac{\ell-\varphi(\ell)}{\ell+\delta}) - G(\ell)\ell\right)^2}.$$

As long as $0 < L_0 < \bar{L}$, we conclude that $\partial \varphi(L)/\partial \delta < 0$ in a right-neighborhood of L_0, which is unsurprising as the threshold loss L_0 for an attack is bound to increase. For larger values of $L > L_0$ the effect may go in opposite directions, so that it is possible that φ increases in delta for *some* losses. The monotonicity properties of $p(L)$ and $q(L)$ with respect to δ are identical. Since $\varphi(L) = L - p(L)(L+\delta)$, it is

$$\frac{\partial p(L)}{\partial \delta} > 0 \;\Leftrightarrow\; \frac{\partial \varphi(L)}{\partial \delta} > -p(\delta) \;\Leftrightarrow\; \frac{\partial \ln(L - \varphi(L))/\partial \delta}{\partial \ln(L+\delta)/\partial \delta} > -1.$$

Concrete instances of the generic nonmonotonicity of $(p(L), q(L))$ in the deterrent are provided in an example in the main text, which concludes our proof. ■

Proof of Corollary 4. If $L_0 \geq \bar{L}$, there is nothing to prove. Assume that $L_0 = \delta\rho/(1-\rho) < \bar{L}$. In the proof of Corollary 2 we have shown that as the number of targets K increases, the agent's payoff $\varphi(L; K)$ *strictly* increases for $L > L_0$. Thus, the security-investment equilibrium $(p(L;K), q(L;K))$ decreases in K for $L > L_0$. Thus, for any fixed $L \in (L_0, \bar{L})$, the sequence $\{p(L;K)\}_{K=2}^{\infty}$ is decreasing and bounded from below by ρ. It therefore converges, and there exists $\hat{\rho} = \hat{\rho}(L) \in [\rho, 1]$ such that $\lim_{K\to\infty} p(L;K) = \hat{\rho}$. Suppose that $\hat{\rho} > \rho$. Then target 1's value of his payoff-maximization problem must be such that

$$\max_{\varphi_1}\left\{-F^{K-1}(\varphi^{-1}(\varphi_1;K))\frac{\varphi_1+\delta}{L_1+\delta}L_1 - \hat{C}(\varphi_1, L_1)\right\}$$

$$\to -F^{K-1}(L_1)(1-\hat{\rho})L_1 - C(\hat{\rho}) \to -C(\hat{\rho}),$$

as $K \to \infty$, where $L_1 = L \in (L_0, \bar{L})$, because $\lim_{K\to\infty} F^{K-1}(L_1) = 0$. Thus, for large K it is best for the agent to reduce p to a value smaller than $\hat{\rho}$, which is feasible since by assumption $\hat{\rho} > \rho$. But this is a contradiction to the monotonicity of the sequence and to $\hat{\rho}$ being the limit, which completes our proof.■

Proof of Proposition 5. Given the regulator's decision variables δ and ρ, let $(p^*(L;\delta,\rho), q^*(L;\delta,\rho))$ be a security-investment equilibrium. Ex-ante social welfare is of the form

$$W(\delta,\rho) = K\int_0^{\bar{L}} \Pi^*(\ell;\delta,\rho)dF(\ell) - D(\delta),$$

or equivalently,

$$W(\delta,\rho) = -KC(\rho)F(L_0) + K\int_{L_0}^{\bar{L}} \Pi^*(\ell;\delta,\rho)dF(\ell) - D(\delta),$$

where any firm's expected equilibrium payoff as a function of its loss L is given by

$$\Pi^*(L;\delta,\rho) = -\mathbf{1}_{\{L\geq L_0\}}G(L)(1-p^*(L;\delta,\rho))L - C(p^*(L;\delta,\rho)).$$

Hence,

$$\begin{aligned} \frac{1}{K}\frac{\partial W(\delta,\rho)}{\partial\rho} \quad &= \quad -C'(\rho)F(L_0) - C(\rho)f(L_0)\frac{\partial L_0}{\partial\rho} \\ &\quad -\Pi^*(L_0;\delta,\rho)f(L_0)\frac{\partial L_0}{\partial\rho} \\ &\quad +\int_{L_0}^{\bar{L}}\frac{\partial\Pi^*(\ell;\delta,\rho)}{\partial\rho}dF(\ell), \end{aligned}$$

where

$$\Pi^*(L_0;\delta,\rho) = -G(L_0)(1-\rho)L_0 - C(\rho) = -G(L_0)\delta\rho - C(\rho).$$

Using the fact that $\partial L_0/\partial\rho = 1/(1-\rho)^2$, we therefore find

$$\frac{\partial W(\delta,\rho)}{\partial\rho} = \frac{\delta\rho\, G(L_0)f(L_0)}{(1-\rho)^2} - C'(\rho)F(L_0) - \int_{L_0}^{\bar{L}}\lambda(\ell)dF(\ell).$$

By virtue of the envelope theorem we have that

$$\frac{\partial\Pi^*(L;\delta,\rho)}{\partial\rho} = \left.\frac{\partial\mathcal{L}}{\partial\rho}\right|_{(p^*(L;\delta,\rho);\delta,\rho)} = -\lambda(L;\delta,\rho) \leq 0,$$

where Lagrange multiplier $\lambda(L;\delta,\rho)$ associated with the inequality constraint $p^*(L;\delta,\rho)\geq\rho$ is

$$\begin{aligned} \lambda(L;\delta,\rho) \quad &= \quad \left[C'(p^*(L;\delta,\rho)) - G(L)L - (L+\delta)g(L)L\right]_+ \\ &\geq \quad 0. \end{aligned} \tag{26}$$

This implies Eq. (11) as the optimality condition for the socially optimal safety standard $\hat{\rho}$. Consider now

$$\begin{aligned} \frac{1}{K}\frac{\partial W(\delta,\rho)}{\partial\delta} \quad &= \quad -C(\rho)f(L_0)\frac{\partial L_0}{\partial\delta} - \Pi^*(L_0;\delta,\rho)f(L_0)\frac{\partial L_0}{\partial\delta} \\ &\quad + \int_{L_0}^{\bar{L}}\frac{\partial\Pi^*(\ell;\delta,\rho)}{\partial\delta}dF(\ell) - \frac{D'(\delta)}{K}, \end{aligned} \tag{27}$$

or equivalently,

$$\frac{1}{K}\frac{\partial W(\delta,\rho)}{\partial\delta} = \frac{\delta\rho^2 G(L_0)f(L_0)}{1-\rho} + \int_{L_0}^{\bar{L}}\frac{\partial\Pi^*(\ell;\delta,\rho)}{\partial\delta}dF(\ell) - \frac{D'(\delta)}{K}.$$

By the envelope theorem we have that

$$\begin{aligned} \frac{\partial\Pi^*(L;\delta,\rho)}{\partial\delta} \quad &= \quad \left.\frac{\partial\mathcal{L}}{\partial\delta}\right|_{(p^*(L;\delta,\rho);\delta,\rho)} \\ &= \quad -\frac{G(L)L + \lambda(L;\delta,\rho)}{L+\delta}\, p^*(L;\delta,\rho) \leq 0, \end{aligned}$$

where the Lagrange multiplier $\lambda(L;\delta,\rho)$ is as in Eq. (26). Thus, the first-order condition for the optimal deterrent $\hat{\delta}$ is given by (12), which concludes our proof. ■